Justin Sherwood - May 6, 2024 – Final project summary.

Final Project - using Python to perform an analysis of patient heart disease data. The potential client for this analysis could be a healthcare provider, such as a hospital, clinic, or health insurance company, looking to better understand and address heart disease risk factors among their patient population

Dataset used - <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>

[downloaded the cleaned dataset - Indicators of Heart Disease](https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease)

Tools used: Anaconda Jupyter Notebook, CMP262 coursework and textbook. Used <https://www.w3schools.com/python> for syntax, and https://stackoverflow.com/.

Python packages used: Pandas for data ingestion, matplotlib and seaborn for visualization, numpy and scipy.stats for statistical analysis.

Analysis goal:

1. What is the correlation between BMI and the likelihood of having heart disease?
2. How does smoking impact the prevalence of heart disease among different age groups?
3. Is there a significant difference in the occurrence of heart disease between males and females?
4. What is the association between alcohol consumption and the incidence of heart disease?
5. Are there any notable disparities in heart disease rates among different racial groups?

Insights Discovered:

* Question 1 analysis reflected a trend of increasing proportions of heart disease cases from healthy weight to overweight and obesity categories, with the most significant proportion observed in the obesity category.
* Question 2 analysis suggested that there are significant differences between the age categories being measured, and the count of patients with heart disease increased as individuals move into older age groups. The findings reinforce the well-established link between smoking and cardiovascular disease.
* Question 5 analysis showed both the visual representation and the statistical test results concluded that there are significant disparities in heart disease rates across the racial groups included in the analysis.

Client Recommendations

* By examining the correlation between various factors and heart disease, the client can identify demographic groups or individuals with a higher risk of developing heart disease. This information can help target preventive interventions more effectively.
* By examining lifestyle factors such as smoking, alcohol consumption, and demographics influence heart disease risk allows the client to develop targeted intervention strategies
* Hospitals can use the findings as a basis for further investigation into the mechanisms underlying heart disease development and potential avenues for treatment or prevention.